JOURNAL OF MULTIVARIATE ANALYSIS 3, 236–242 (1973)

# An Axiomatic Foundation for a Multivariate Measure of Affinity among a Number of Distributions*

H. KAUFMAN AND A. M. MATHAI

*Department of Mathematics, McGill University, Montreal, Quebec, Canada*

*Communicated by A. P. Dempster*

The aim of this article is to give an axiomatic characterization of the multivariate measure of "affinity" among a number of distributions. This measure has been used in the last few years for statistical decision making. When there are only two distributions involved this measure of affinity is related to a measure of distance between the distributions. In this article the multivariate measure is characterized by a set of four axioms. It is shown that these axioms lead to a functional equation whose unique solution is the measure of affinity under consideration. The axioms are also intuitively appealing as desirable properties for a measure of affinity.

## 1. INTRODUCTION

Consider a set of $r$ discrete distributions $(P_{i1}, P_{i2}, \ldots, P_{in})$, for $i = 1, 2, \ldots, r$, where $P_{ij} \geqslant 0$, for all $i, j$, $\sum_{j=1}^{n} P_{ij} = 1$, for $i = 1, 2, \ldots, r$. Matusita (1967) considered a measure which would in some sense indicate the closeness of the distributions to one another which he called a measure of affinity among the distributions. The measure so defined is as follows:

$$\rho_n = \sum_{j=1}^{n} (P_{1j} P_{2j} \cdots P_{rj})^{1/r}, \qquad 2 \leqslant r < \infty. \tag{1.1}$$

When $r = 2$, it is easily seen that $\rho_n$ is related to a distance measure as follows:

$$D_n = 2(1 - \rho_n), \tag{1.2}$$

Received December 6, 1971; revised October 26, 1972.

AMS 1970 subject classifications: Primary 62H30.

Key words: Axiomatic definition, characterization, measure of affinity, axiomatic foundations.

* This research is supported by National Research Council of Canada Grant No. A 3057.

where

$$D_n = \sum_{j=1}^{n} (P_{1j}^{1/2} - P_{2j}^{1/2})^2. \tag{1.3}$$

$D_n$ can be interpreted as the square of a measure of distance between the two populations $(P_{i1}, P_{i2}, \ldots, P_{in})$, $i = 1, 2$ and, thus, $\rho_n$ can be given a geometrical interpretation as $\cos\theta$, where $\theta$ is the angle between the vectors if the two distributions are represented as points on a hypersphere. When $r = 2$, $\rho_n$ is used extensively in statistical inference problems, some of which may be seen from Matusita (1966). In the general case some of its properties are studied and some applications are pointed out in Kirmani (1968, 1971) and Matusita (1967, 1971).

The aim of this article is to give an axiomatic foundation for $\rho_n$ in (1.1) for a fixed $r$, $2 \leqslant r < \infty$. The main result is embodied in the following theorem.

## 2. Characterization of Affinity Measure

THEOREM. *Let $A_n(P_{11}, P_{12}, \ldots, P_{1n}; P_{21}, \ldots, P_{2n}; \ldots; P_{r1}, P_{r2}, \ldots, P_{rn})$, $(n \geqslant 2)$ be a function of $P_{ij}$, $P_{ij} \geqslant 0$ for all $i = 1, \ldots, r$, $j = 1, 2, \ldots, n$ and $\sum_{j=1}^{n} P_{ij} = 1$ for $i = 1, 2, \ldots, r$, satisfying the following properties $R_1$, $R_2$, $R_3$, and $R_4$.*

$R_1$: *Recursivity.*

$$\begin{aligned} &A_n(P_{11}, \ldots, P_{1n}; P_{21}, \ldots, P_{2n}; \ldots; P_{r1}, \ldots, P_{rn}) \\ &\quad = A_{n-1}(P_{11} + P_{12}, P_{13}, \ldots, P_{1n}; P_{21} + P_{22}, P_{23}, \ldots, P_{2n}; \ldots; P_{r1} \\ &\qquad + P_{r2}, P_{r3}, \ldots, P_{rn}) + [(P_{11} + P_{12})(P_{21} + P_{22}) \cdots (P_{r1} + P_{r2})]^{1/r} \\ &\qquad \times \left[A_2\left(\frac{P_{11}}{P_{11} + P_{12}}, \frac{P_{12}}{P_{11} + P_{12}}; \ldots; \frac{P_{r1}}{P_{r1} + P_{r2}}, \frac{P_{r2}}{P_{r1} + P_{r2}}\right) - 1\right], \end{aligned}$$

*for $P_{i1} + P_{i2} > 0$, $i = 1, \ldots, r$; $2 \leqslant r < \infty$ and $n \geqslant 3$.*

$R_2$: *Symmetry. $A_3$ is symmetric in r-tuples*

$$(P_{1j}, P_{2j}, \ldots, P_{rj}) \qquad j = 1, 2, 3.$$

$R_3$: *Normalization.*

$$A_2\left(\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2}; \ldots; \frac{1}{2}, \frac{1}{2}; \frac{1}{4}, \frac{3}{4}\right) = \frac{1}{2}\left[\frac{3^{1/r} + 1}{2^{1/r}}\right]$$

$R_4$: *$A_2(x_1, 1 - x_1; x_2, 1 - x_2; \ldots; x_r, 1 - x_r)$ is continuous in its variables $x_1, \ldots, x_r \in I = [0, 1]$.*

*Then $R_1$, $R_2$, $R_3$ and $R_4$ uniquely determine $A_n$ as Matusita's measure of "affinity" as given on the right side of* (1.1).

*Note.* $R_1$ and $R_2$ will determine $A_n$ as a constant multiple of the right side of (1.1) and $R_3$ and $R_4$ are used to evaluate this constant. Conversely, it is easily seen that the $\rho_n$ defined in (1.1) satisfies properties $R_1$ to $R_4$.

*Proof.* Let

$$h(x_1, ..., x_r) = A_2(x_1, 1 - x_1; x_2, 1 - x_2; ...; x_r, 1 - x_r) - 1, \quad \text{for} \quad x_1, ..., x_r \in I. \tag{2.1}$$

It can be shown that,

$$h(x_1, ..., x_r) = h(1 - x_1, ..., 1 - x_r) \qquad \text{for} \quad x_1, ..., x_r \in I, \tag{2.2}$$

and that $h(x_1, ..., x_r)$ satisfies the functional equation

$$\begin{aligned} h(x_1, ..., x_r) &+ [(1 - x_1)(1 - x_2) \cdots (1 - x_r)]^{1/r} h\left(\frac{u_1}{1 - x_1}, \frac{u_2}{1 - x_2}, ..., \frac{u_2}{1 - x_r}\right) \\ &= h(u_1, ..., u_r) + [(1 - u_1) \cdots (1 - u_r)]^{1/r} h\left(\frac{x_1}{1 - u_1}, \frac{x_2}{1 - u_2}, ..., \frac{x_r}{1 - u_r}\right), \\ &\text{for} \quad x_1, ..., x_r, \quad u_1, ..., u_r \in [0, 1), \quad x_1 + u_1, ..., x_r + u_r \in I. \end{aligned} \tag{2.3}$$

The proof of the theorem will be established with the help of the following lemma.

Lemma. *$h(x_1, ..., x_r)$, satisfying the functional equation* (2.3), *is uniquely determined as*

$$h(x_1, ..., x_r) = (x_1 x_2 \cdots x_r)^{1/r} + [(1 - x_1)(1 - x_2) \cdots (1 - x_r)]^{1/r} - 1, \quad \textit{for} \quad x_1, ..., x_r \in I. \tag{2.4}$$

*Proof.* Consider the function

$$\begin{aligned} g(x_1&, ..., x_r; u_1, ..., u_r) \\ &= h(x_1, ..., x_r) + [(x_1 \cdots x_r)^{1/r} + \{(1 - x_1) \cdots (1 - x_r)\}^{1/r}]\, h(u_1, u_2, ..., u_r), \\ &\text{for} \quad x_1, ..., x_r, u_1, ..., u_r \in (0, 1). \end{aligned} \tag{2.5}$$

We will show that $g(x_1, ..., x_r; u_1, ..., u_r)$ is symmetric in $(x_1, u_1), ..., (x_r, u_r)$.

Setting $u_1/(1 - x_1) = v_1, \dots, u_r/(1 - x_r) = v_r$ ; $1 - x_1 = y_1, \dots, 1 - x_r = y_r$ in (2.3) and using (2.2) we have

$$h(y_1, \dots, y_r) + (y_1 \cdots y_r)^{1/r} h(v_1, \dots, v_r)$$
$$= h(y_1 v_1, \dots, y_r v_r) + [(1 - y_1 v_1) \cdots (1 - y_r v_r)]^{1/r} h\left(\frac{1 - y_1}{1 - y_1 v_1}, \dots, \frac{1 - y_r}{1 - y_r v_r}\right),$$
$$y_i \neq 0, \qquad y_i v_i \neq 1, \qquad i = 1, \dots, r. \quad (2.6)$$

Interchanging the $y$'s and $v$'s in (2.6) we get

$$h(v_1, \dots, v_r) + (v_1 \cdots v_r)^{1/r} h(y_1, \dots, y_r)$$
$$= h(y_1 v_1, \dots, y_r v_r) + [(1 - y_1 v_1) \cdots (1 - y_r v_r)]^{1/r} h\left(\frac{1 - v_1}{1 - y_1 v_1}, \dots, \frac{1 - v_r}{1 - y_r v_r}\right),$$
$$v_i \neq 0, \qquad y_i v_i \neq 1, \qquad i = 1, \dots, r. \quad (2.7)$$

Now by eliminating $h(y_1 v_1, \dots, y_r v_r)$ we get

$$h(y_1, \dots, y_r) + (y_1 \cdots y_r)^{1/r} h(v_1, \dots, v_r)$$
$$= h(v_1, \dots, v_r) + (v_1 \cdots v_r)^{1/r} h(y_1, \dots, y_r) + [(1 - y_1 v_1) \cdots (1 - y_r v_r)]^{1/r}$$
$$\times \left[h\left(\frac{1 - y_1}{1 - y_1 v_1}, \dots, \frac{1 - y_r}{1 - y_r v_r}\right) - h\left(\frac{1 - v_1}{1 - y_1 v_1}, \dots, \frac{1 - v_r}{1 - y_r v_r}\right)\right]. \quad (2.8)$$

Replacing $y_i$ by $(1 - y_i)/(1 - y_i v_i)$, $i = 1, \dots, r$, in (2.6) we get

$$h\left(\frac{1 - y_1}{1 - y_1 v_1}, \dots, \frac{1 - y_r}{1 - y_r v_r}\right) + \left[\left(\frac{1 - y_1}{1 - y_1 v_1}\right) \cdots \left(\frac{1 - y_r}{1 - y_r v_r}\right)\right]^{1/r} h(v_1, \dots, v_r)$$
$$= h\left[\frac{(1 - y_1) v_1}{1 - y_1 v_1}, \dots, \frac{(1 - y_r) v_r}{1 - y_r v_r}\right]$$
$$+ \left[\left(\frac{1 - v_1}{1 - y_1 v_1}\right) \cdots \left(\frac{1 - v_r}{1 - y_r v_r}\right)\right]^{1/r} h(y_1, \dots, y_r),$$

by using (2.2). Thus,

$$[(1 - y_1 v_1) \cdots (1 - y_r v_r)]^{1/r} \left[h\left(\frac{1 - y_1}{1 - y_1 v_1}, \dots, \frac{1 - y_r}{1 - y_r v_r}\right) - h\left(\frac{1 - v_1}{1 - y_1 v_1}, \dots, \frac{1 - v_r}{1 - y_r v_r}\right)\right]$$
$$= [(1 - v_1) \cdots (1 - v_r)]^{1/r} h(y_1, \dots, y_r) - [(1 - y_1)(1 - y_2) \cdots (1 - y_r)]^{1/r} h(v_1, \dots, v_r).$$

Substituting in (2.8) gives

$$\begin{aligned} h(y_1,\ldots,y_r) &+ (y_1\cdots y_r)^{1/r}h(\nu_1,\ldots,\nu_r) + [(1-y_1)\cdots(1-y_r)]^{1/r}h(\nu_1,\ldots,\nu_r) \\ &= h(\nu_1,\ldots,\nu_r) + (\nu_1\cdots\nu_r)^{1/r}h(y_1,\ldots,y_r) \\ &\quad + [(1-\nu_1)\cdots(1-\nu_r)]^{1/r}h(y_1,\ldots,y_r). \end{aligned} \tag{2.9}$$

Thus,

$$h(y_1,\ldots,y_r) = \frac{[(y_1\cdots y_r)^{1/r} + \{(1-y_1)\cdots(1-y_r)\}^{1/r} - 1]\, h(\nu_1,\ldots,\nu_r)}{[(\nu_1\cdots\nu_r)^{1/r} + \{(1-\nu_1)\cdots(1-\nu_r)\}^{1/r} - 1]}. \tag{2.10}$$

Since arbitrary values can be assigned for $\nu_1,\ldots,\nu_r$, we have, using $R_3$,

$$h(y_1,\ldots,y_r) = (y_1\cdots y_r)^{1/r} + \{(1-y_1)\cdots(1-y_r)\}^{1/r} - 1, \qquad \text{for } y_1,\ldots,y_r \in (0,1). \tag{2.11}$$

From (2.11) we also observe that $A_2(y_1, 1-y_1; \ldots; y_r, 1-y_r) = 1$ when $y_1 = \cdots = y_r$. This implies that affinity is maximum when the distributions are identical. Also it follows from (2.6) by putting $y_i = 1$, $i = 1, 2,\ldots, r$ and then using (2.2) that

$$h(1,\ldots,1) = h(0,\ldots,0) = 0. \tag{2.12}$$

We now extend (2.11) to the closed interval $I = [0, 1]$. In (2.3) put $x_1 = x_2 = \cdots = x_k = 0$, $u_{k+1} = \cdots = u_r = 0$ $(k < r)$. By rearranging the terms we get

$$\begin{aligned} h(0,\ldots,0,x_{k+1},\ldots,x_r)&\{[(1-u_1)\cdots(1-u_k)]^{1/r} - 1\} \\ &= h(u_1,\ldots,u_k,0,\ldots,0)\{[(1-x_{k+1})\cdots(1-x_r)]^{1/r} - 1\}. \end{aligned} \tag{2.13}$$

Since the $x$'s and $u$'s are arbitrary in $[0, 1)$, from (2.17) we obtain

$$h(0,\ldots,0,x_{k+1},\ldots,x_r) = c_1\{[(1-x_{k+1})\cdots(1-x_r)]^{1/r} - 1\}, \qquad x_{k+1},\ldots,x_r \in [0,1), \tag{2.14}$$

where $c_1$ is a constant, not involving $x_{k+1},\ldots,x_r$. But $h(0,\ldots,0,x_{k+1},\ldots,x_r) = h(1,\ldots,1,1-x_{k+1},\ldots,1-x_r)$ according to (2.2), and, therefore, by replacing $x$'s by $1-x$'s we get

$$h(1,\ldots,1,x_{k+1},\ldots,x_r) = c_2\{(x_{k+1}\cdots x_r)^{1/r} - 1\},\ x_{k+1},\ldots,x_r \in (0,1]. \tag{2.15}$$

From the derivation of (2.14) and (2.15) it is evident that instead of $x_1 = \cdots = x_k = 0$ any of the $x$'s could be taken as zeros in (2.14) and similarly any of the $x$'s could be taken as unities in (2.14). In (2.3) $x_i + u_i$, $i = 1, 2,\ldots, r$, can take value unity. Putting $u_1 = 1 - x_1$, $x_1 \in (0, 1]$, $x_2 = 0 = \cdots = x_k$, $u_{k+1} = 0 = \cdots = u_r$, proceeding in a similar fashion and finally evaluating at $x_1 = 1$, we get

$$h(1, 0, \ldots, 0, x_{k+1}, \ldots, x_r) = c_2(-1), \tag{2.16}$$

where $c_2$ is a constant. Similar results are easily obtained for any number of unities and zeros in $h(\ )$. Hence, in general, using $R_3$, $R_4$ and (2.11) we have

$$h(x_1, \ldots, x_r) = (x_1 \cdots x_r)^{1/r} + [(1 - x_1) \cdots (1 - x_r)]^{1/r} - 1, \quad \text{for} \quad x_1, \ldots, x_r \in I. \tag{2.17}$$

Thus, we obtain

$$\begin{aligned} &A_2(x_1, 1 - x_1; x_2, 1 - x_2; \ldots; x_r, 1 - x_r) \\ &\quad = (x_1 \cdots x_r)^{1/r} + [(1 - x_1) \cdots (1 - x_r)]^{1/r}, \qquad \text{for} \quad x_1, \ldots, x_r \in I. \end{aligned} \tag{2.18}$$

By repeated applications of the recursivity property we obtain

$$\begin{aligned} &A_n(P_{11}, P_{12}, \ldots, P_{1n}; P_{21}, P_{22}, \ldots, P_{2n}; \ldots; P_{r1}, P_{r2}, \ldots, P_{rn}) - 1 \\ &\quad = \sum_{i=2}^{n} (\mathbf{P}_{1i}\mathbf{P}_{2i} \cdots \mathbf{P}_{ri})^{1/r} \\ &\quad\quad \times \left\{ A_2 \left( \frac{P_{1i}}{\mathbf{P}_{1i}}, 1 - \frac{P_{1i}}{\mathbf{P}_{1i}}; \frac{P_{2i}}{\mathbf{P}_{2i}}, 1 - \frac{P_{2i}}{\mathbf{P}_{2i}}; \ldots; \frac{P_{ri}}{\mathbf{P}_{ri}}, 1 - \frac{P_{ri}}{\mathbf{P}_{ri}} \right) - 1 \right\}, \end{aligned} \tag{2.19}$$

where

$$\mathbf{P}_{ji} = P_{j1} + P_{j2} + \cdots + P_{ji}, \qquad j = 1, 2, \ldots, r. \tag{2.20}$$

Substitution of $A_2$ from (2.18) yields

$$\begin{aligned} A_n - 1 &= \sum_{i=2}^{n} (\mathbf{P}_{1i} \cdots \mathbf{P}_{ri})^{1/r} \\ &\quad \times \left\{ \left( \frac{P_{1i} \cdots P_{ri}}{\mathbf{P}_{1i} \cdots \mathbf{P}_{ri}} \right)^{1/r} + \left[ \left(1 - \frac{P_{1i}}{\mathbf{P}_{1i}}\right) \cdots \left(1 - \frac{P_{ri}}{\mathbf{P}_{ri}}\right) \right]^{1/r} - 1 \right\} \\ &= \sum_{i=2}^{n} (P_{1i}P_{2i} \cdots P_{ri})^{1/r} + \sum_{i=2}^{n} [(\mathbf{P}_{1i} - P_{1i}) \cdots (\mathbf{P}_{ri} - P_{ri})]^{1/r} \\ &\quad - \sum_{i=2}^{n} (\mathbf{P}_{1i} \cdots \mathbf{P}_{ri})^{1/r}. \end{aligned} \tag{2.21}$$

But

$$\mathbf{P}_{ji} - P_{ji} = \mathbf{P}_{j,i-1}, \qquad j = 1, 2, \ldots, r, \tag{2.22}$$

and, thus,

$$\sum_{i=2}^{n} (\mathbf{P}_{1,i-1} \cdots \mathbf{P}_{r,i-1})^{1/r} - \sum_{i=2}^{n} (\mathbf{P}_{1i} \cdots \mathbf{P}_{ri})^{1/r}$$
$$= (\mathbf{P}_{11}\mathbf{P}_{21} \cdots \mathbf{P}_{r1})^{1/r} - (P_{1n} \cdots P_{rn})^{1/r} = (P_{11}P_{21} \cdots P_{r1})^{1/r} - 1. \quad (2.23)$$

That is,

$$A_n(P_{11}, P_{12}, ..., P_{1n}; P_{21}, ..., P_{2n}; ...; P_{r1}, ..., P_{rn}) = \sum_{i=1}^{n} (P_{1i}P_{2i} \cdots P_{ri})^{1/r}. \quad (2.24)$$

This completes the proof that Matusita's measure of affinity among $r$ discrete distributions is uniquely determined by properties $R_1$, $R_2$, $R_3$, and $R_4$.

## References

[1] Kirmani, S. N. U. A. (1968). Some results on Matusita's measure of distance. *J. Indian Statist. Assoc.* **6** 89–98.

[2] Kirmani, S. N. U. A. (1971). Some limiting properties of Matusita's measure of distance. *Ann. Inst. Statist. Math.* **23** 157–162.

[3] Matusita, K. (1966). A distance and related statistic in multivariate analysis. In *Multivariate Analysis* (P. R. Krishnaiah, Ed.), pp. 187–200. Academic Press, New York.

[4] Matusita, K. (1967). On the notion of affinity of several distributions and some of its applications. *Ann. Inst. Statist. Math.* **19** 181–192.

[5] Matusita, K. (1971). Some properties of affinity and applications. *Ann. Inst. Statist. Math.* **23** 137–155.